高々スター次数nの拡張正規表現(アルゴリズムと計算量の理論)

劉, 信根; 橋口, 攻三郎

数理解析研究所講究録
731巻 1990年 143-154

高々スター次数nの拡張正規表現

Extended Regular Expressions of Arbitrary Star Degrees

劉 僖根　　橋口 攻三郎
Heekeun YOO and Kosaburo HASHIGUCHI
豊橋技術科学大学
Toyohashi University of Technology

1. まえかき

[7, 13]において，われわれは次数2のスター演算と次数(2,1)のスター演算の概念を導入し，アルファベットΣ上の高々スター次数2の拡張正規言語のクラス(ERL(2,Σ))と高々スター次数(2,1)の拡張AL正規言語のクラス(EARL(2,1,Σ))のいくつかの性質を明らかにした．

本研究では，次数n $(n \geq 3)$のスター演算の概念を導入し，和集合，連接，Kleene-閉包，高々次数n $(n \geq 3)$のスター演算に関する有限言語族の閉包，すなわち，アルファベットΣ上の高々スター次数nの拡張正規言語のクラス(ERL(n,Σ))のいくつかの性質を明らかにする．特に，(1) ERL(n,Σ) (n=2, 3, 4, …)の階層構造は無限である，(2) $n \geq 3$に対して，ERL(n,Σ)は文脈依存言語の部分族である，(3) ERL(3,Σ) - CFL(Σ) $\neq \phi$，(4)任意の$n \geq 3$に対して，LCFL(Σ) - ERL(n,Σ) $\neq \phi$，等を示す．

2. 準備

Σは空でない有限アルファベットである．λは空系列である．φは空集合である．任意の$W \in \Sigma^*$に対して$l(W)$はWの長さで，W^RはWの逆系列である．$L \subset \Sigma^*$に対して$L^R=\{W^R | W \in L\}$である．任意のアルファベットV, V′と$W \in V^*$に対して$V'(W)$はWに現われるV′の記号の集合である．任意の$A \in V$に対して$\#_A(W)$はWに現われるAの個数である．集合Bに対して#BはBの濃度である．

［定義2．1］

任意の$n \geq 2$に対して，ERE(n,Σ)はΣ上の高々スター次数nの拡張正規表現のクラスであり，次のように帰納的に定義される．

(1) λ, φ, $a \in$ ERE(n,Σ)（但し$a \in \Sigma$）；

(2) もし$E_1, E_2 \in$ ERE(n,Σ)であれば，$E_1 \cup E_2$, E_1E_2, $(E_1)^* \in$ ERE(n,Σ)である；

(3) 任意の$2 \leq m \leq n$, $p \geq 1$に対して，$E_{11}, \cdots, E_{1p}, E_{01}, E_{21}, \cdots, E_{2p}, E_{02}, \cdots, E_{0m-1}, E_{m1}, \cdots, E_{mp} \in$ ERE(n,Σ)であれば，$(E_{11}, \cdots, E_{1p}, E_{01}, E_{21}, \cdots, E_{2p}, E_{02}, \cdots, E_{0m-1}, E_{m1}, \cdots, E_{mp})^{m*} \in$ ERE(n,Σ)である；

［注意2．1］

(1) 各々の$E \in$ ERE(n,Σ)に対して，不必要なかっこはしばしば省略される．

(2) 表現を明瞭にするため$(E_{11}, \cdots, E_{1p}, E_{01}, E_{21}, \cdots, E_{2p}, E_{02}, \cdots, E_{0n-1}, E_{n1}, \cdots, E_{np})^{n*}$

の代わりにしばしば

$$\left[\begin{bmatrix}E_{11}\\ \vdots\\ E_{1p}\end{bmatrix}, E_{01}, \begin{bmatrix}E_{21}\\ \vdots\\ E_{2p}\end{bmatrix}, E_{02}, \cdots, E_{0n-1}, \begin{bmatrix}E_{n1}\\ \vdots\\ E_{np}\end{bmatrix}\right]^{n*}$$ の表現が導入される.

[定義2. 2]

$n \geq 2$ に対して, $E \in ERE(n,\Sigma)$ によって表現される言語 $|E|$ は次のように帰納的に定義される.

(1) $|\lambda|=\{\lambda\}$, $|\phi|=\phi$, $|a|=\{a\}$ (但し $a \in \Sigma$) ;

(2) $|E_1 \cup E_2|=|E_1| \cup |E_2|$;

$|E_1E_2|=|E_1||E_2|=\{vw \mid v \in |E_1|, w \in |E_2|\}$;

$|(E)^*|=|E|^*=\bigcup_{i=0}^{\infty} |E|^i$;

(3) $|(E_{11}, \cdots, E_{1p}, E_{01}, E_{21}, \cdots, E_{2p}, E_{02}, \cdots, E_{0m-1}, E_{m1}, \cdots, E_{mp})^{m*}|$

$=|E_{01}E_{02} \cdots E_{0m-1}| \cup (\bigcup_{i=1}^{\infty} (\bigcup_{1 \leq j_1, \cdots, j_i \leq p} L_1|E_{01}|L_2|E_{02}| \cdots L_{m-1}|E_{0m-1}|L_m))$,

ここで, $1 \leq k \leq m$ に対して, もし k が奇数であれば, $L_k=|E_{kj_1}||E_{kj_2}| \cdots |E_{kj_i}|$ であり, 偶数であれば, $L_k=|E_{kj_i}||E_{kj_{i-1}}| \cdots |E_{kj_1}|$ である.

$E \in ERE(n,\Sigma)$ $(n \geq 2)$ に対して, $|E|$ は Σ 上の高々スター次数 n の拡張正規言語と呼ばれる.

[定義2. 3]

$n \geq 2$ に対して, $ERL(n,\Sigma)$ は Σ 上の高々スター次数 n の拡張正規言語のクラスである.

[例2. 1]

(1) $|(a, \lambda, b, \lambda, c)^{3*}|=\{a^nb^nc^n \mid n \geq 0\}$;

(2) $\left|\left[\begin{bmatrix}a\\ b\end{bmatrix}, c, \begin{bmatrix}a\\ b\end{bmatrix}, c, \begin{bmatrix}a\\ b\end{bmatrix}, c, \begin{bmatrix}a\\ b\end{bmatrix}\right]^{4*}\right|=\{wcw^Rcwcw^R \mid w \in \{a,b\}^*\}$.

(3) 任意の $n \geq 3$, $2 \leq m < n$, $p \geq 1$, $E=(E_{11}, \cdots, E_{1p}, E_{01}, E_{21}, \cdots, E_{2p}, E_{02}, \cdots, E_{0m-1}, E_{m1}, \cdots, E_{mp})^{m*} \in ERL(n,\Sigma)$ に対して, 次が成立する:

$|E|=|(E_{11}, \cdots, E_{1p}, E_{01}, E_{21}, \cdots, E_{2p}, E_{02}, \cdots, E_{0m-1}, E_{m1}, \cdots, E_{mp}, F_{0m}, F_{m+11}, \cdots, F_{m+1p},$
$F_{0m+1}, \cdots, F_{0n-1}, F_{n1}, \cdots, F_{np})^{n*}|$,

ここで, すべての $m \leq i \leq n-1$, $m+1 \leq j \leq n$, $1 \leq k \leq p$ に対して, $F_{0i}=F_{jk}=\lambda$ である.

3. ERL(n,Σ) の性質

まず, $ERL(1,\Sigma) \subset ERL(2,\Sigma) \subset ERL(3,\Sigma) \subset \cdots$ の階層構造は $\#\Sigma \geq 2$ のとき無限であることを示す. ここで, Kleene－閉包は次数1のスター演算とみなされる.

[定義3. 1]

$n \geq 2$ に対して, $L_n=|(a_1, \lambda, a_2, \lambda, \cdots, \lambda, a_n)^{n*}|$ である. ここで, もし i が奇数であれば, $a_i=a$ であり, 偶数であれば, $a_i=b$ である (但し $\Sigma=\{a,b\}$).

[補題３．１]

任意の$n \geq m \geq 2$, $u,v,x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_{m-1} \in \Sigma^*$に対して，もし$|u(x_1,y_1,x_2,y_2,\cdots,y_{m-1},x_m)^{m*}v| \subset L_n$であれば，任意の$1 \leq i \leq m$に対して，$x_i \in a^* \cup b^*$である．

（証明）$|u(x_1,y_1,x_2,y_2,\cdots,y_{m-1},x_m)^{m*}v| \subset L_n$とする．そのとき$ux_1y_1x_2y_2 \cdots y_{m-1}x_mv \in L_n$である．従って，ある$j \geq 0$に対して，$ux_1y_1x_2y_2 \cdots y_{m-1}x_mv = a_1{}^ja_2{}^j \cdots a_n{}^j$である．もしある$1 \leq i \leq m$に対して，$x_i \notin a^* \cup b^*$であれば，$k=2n$に対して，$ux_1{}^ky_1x_2{}^ky_2 \cdots y_{m-1}x_m{}^kv \in L_n$である．これは矛盾である．□

[定理３．１]

次の包含階層構造は無限である．

$ERL(1,\Sigma) \subset ERL(2,\Sigma) \subset ERL(3,\Sigma) \subset \cdots$．（ここで，$\#\Sigma \geq 2$である）．

次の補題を証明すれば十分である．

[補題３．２]

任意の$n \geq 2$に対して，

(1) $ERL(n-1,\Sigma) \subset ERL(n,\Sigma)$; (2) $L_n \in ERL(n,\Sigma)-ERL(n-1,\Sigma)$.

ここで，$\Sigma \supset \{a, b\}$.

（証明）(1) 明らか．(2). $L_n \in ERL(n-1,\Sigma)$を仮定する．明らかに$n \geq 4$である．そのとき$|E|=L_n$であるような$E \in ERE(n-1,\Sigma)$が存在する．そのとき$|u(x_1,y_1,x_2,y_2,\cdots,y_{n-2},x_{n-1})^{(n-1)*}v| \subset |E|=L_n$であるような$u,v,x_1,x_2,\cdots,x_{n-1},y_1,y_2,\cdots,y_{n-2} \in \Sigma^*$が存在する，ここで$L_n$は無限集合であるから，ある$i\,(1 \leq i \leq n-1)$に対して$x_i \neq \lambda$と仮定してよい．補題3.1によって，$x_i \in a^+ \cup b^+$である．そのとき$ux_1y_1x_2y_2 \cdots y_{n-2}x_{n-1}v \in L_n$である．従って，ある$j \geq 1$に対して，$ux_1y_1x_2y_2 \cdots y_{n-2}x_{n-1}v= a_1{}^ja_2{}^j \cdots a_n{}^j$である．$x_k \in a^* \cup b^*\,(1 \leq k \leq n-1)$であるから，各々の$x_k$は$a_p{}^j\,(1 \leq p \leq n)$の因子である．しかし，$x_k$の数は$n-1$であるから，ある$p\,(1 \leq p \leq n)$に対して$a_p{}^j$は任意の$x_k$を因子として含まない．そのとき
$ux_1{}^2y_1x_2{}^2y_2 \cdots y_{n-2}x_{n-1}{}^2v=a_1^{j_1}a_2^{j_2} \cdots a_{p-1}^{j_{p-1}}a_p^{j}a_{p+1}^{j_{p+1}} \cdots a_n^{j_n} \in L_n$かつある$k \neq p$に対して，$j_k \neq j$である．これは矛盾である．□

Σ上の文脈依存文法G (CSG)は(V, Σ, P, S)の４つの組で示される．Vは変数の有限集合，Σは終端記号，Pは生成規則の有限集合，そしてSは開始記号である．文法Gにおける各々の生成規則は：(1) $\alpha \rightarrow \beta$ ($\alpha \in (V \cup \Sigma)^*V(V \cup \Sigma)^*, \beta \in (V \cup \Sigma)^*$かつ$l(\alpha) \leq l(\beta)$) ; (2) $S \rightarrow \lambda$ （但しSはGにおけるいかなる生成規則の右辺にも現われない）．L(G)は文法Gによって生成される文脈依存言語(CSL)である．

[定義３．２]

$CSL(\Sigma)$はΣ上の文脈依存言語のクラスである．

[定理３．２]

任意の$n \geq 3$, $E \in ERE(n,\Sigma)$に対して，$|E|$を生成する文脈依存文法Gを構成することができる．

（証明）証明はEに現われる演算子の個数に関する帰納法による．

基底段階：$|\lambda|$, $|\phi|$, $|a| \in CSL(\Sigma)$ （但し$a \in \Sigma$）.
帰納段階：$E \in ERE(n, \Sigma)$ は次の４つのいずれかの形をしている.
(1) $E=E_1 \cup E_2$,　(2) $E=E_1E_2$,　(3) $E=(E_1)^*$,
(4) $E=(E_{11}, \cdots, E_{1m}, E_{01}, E_{21}, \cdots, E_{2m}, E_{02}, \cdots, E_{0n-1}, E_{n1}, \cdots, E_{nm})^{n*}$.
(1)-(3)の場合，帰納法の仮定により$|E_1|-\{\lambda\}$と$|E_2|-\{\lambda\}$を生成する文脈依存文法$G_1=(V_1, \Sigma, P_1, S_1)$と$G_2=(V_2, \Sigma, P_2, S_2)$が存在する．そのとき$V_1 \cap V_2 = \phi$と仮定してよい.
(1) $E=E_1 \cup E_2$の場合
　次の文法$G=(V,\Sigma,P,S)$はEに対する条件を満たす；
(1.1) $V=V_1 \cup V_2 \cup \{S\}$；
(1.2) $P=P_1 \cup P_2 \cup S \to S_1|S_2 \cup P'$,
ここで，$\lambda \in |E|$であれば，$P'=S \to \lambda$であり，その他は，$P'=\phi$である.
(2) $E=E_1E_2$の場合
　次の文法$G=(V,\Sigma,P,S)$はEに対する条件を満たす；
(2.1) $V=V_1 \cup V_2 \cup \{S\}$；
(2.2) $P=P_1 \cup P_2 \cup S \to S_1S_2 \cup P'$,
ここで，P'：$\lambda \in |E_1| \cap |E_2|$であれば，$S \to S_1|S_2|\lambda$,
　　P'：$\lambda \notin |E_1|$かつ$\lambda \in |E_2|$であれば，$S \to S_1$,
　　P'：$\lambda \in |E_1|$かつ$\lambda \notin |E_2|$であれば，$S \to S_2$,
　　そして$\lambda \notin |E_1 \cup E_2|$であれば，$P'=\phi$である.
(3) $E=(E_1)^*$の場合
　次の文法$G=(V,\Sigma,P,S)$はEに対する条件を満たす；
(3.1) $V=V_1 \cup \{S, S'\}$；
(3.2) $P=P_1 \cup S \to \lambda|S' \cup S' \to S_1|S_1S'$.
(4) $E=(E_{11}, \cdots, E_{1m}, E_{01}, E_{21}, \cdots, E_{2m}, E_{02}, \cdots, E_{0n-1}, E_{n1}, \cdots, E_{nm})^{n*}$の場合
　$G_{ij}=(V_{ij}, \Sigma, P_{ij}, S_{ij})$を$|E_{ij}|-\{\lambda\}$を生成する文脈依存文法とする．（$i=0, 1 \le j \le n-1$, または, $1 \le i \le n$, $1 \le j \le m$），ここで，集合$\{V_{ij} \mid i=0, 1 \le j \le n-1$, または, $1 \le i \le n$, $1 \le j \le m\}$は互いに素であると仮定してよい.
　次の２つの場合を考える.
(4.1) すべてのi, jに対し，$\lambda \notin |E_{ij}|$の場合
　次の文法$G=(V, \Sigma, P, S)$はEに対する条件を満たす；

(4.1.1) $V=(\bigcup_{i=1}^{n} \bigcup_{j=1}^{m} V_{ij}) \cup (\bigcup_{j=1}^{n-1} V_{0j}) \cup \{L_1, \cdots, L_m, R_1, \cdots, R_m, C_1, \cdots C_m, S_1, \cdots, S_{n-1}, S\}$.

(4.1.2) $P=P' \cup (\bigcup_{i=1}^{n} \bigcup_{j=1}^{m} P_{ij}) \cup (\bigcup_{j=1}^{n-1} P_{0j})$,

ここで，P'は次の規則からなっている：
(4.1.3) $n=3$のとき,
　$S \to S_{01}S_{02}|S_{1j}S_{01}S_{2j}S_{02}S_{3j}$ $(1 \le j \le m)$
　$S \to S_{1j}S_1R_j (1 \le j \le m)$
　$S_1 \to S_{1j}S_1C_j|S_{1j}S_{01}S_{2j}S_2S_{3j} (1 \le j \le m)$

$S_2C_j \to S_{2j}S_2S_{3j}(1 \le j \le m)$
$S_2R_j \to S_{2j}S_{02}S_{3j}(1 \le j \le m)$
$D_k \in \{R_k, C_k \mid 1 \le k \le m\}$ に対して,
$S_{ij}D_k \to D_kS_{ij}($ i=0, $1 \le j \le n-1$, または, $1 \le i \le n$, $1 \le j \le m)$.
(4.1.4) $n \ge 4$ かつ偶数のとき,
$S \to S_{01}S_{02} \cdots S_{0(n-1)} | S_{1j}S_{01}S_{2j}S_{02} \cdots S_{0(n-1)}S_{nj}(1 \le j \le m)$
$S \to S_{1j}S_1R_j(1 \le j \le m)$
$S_1 \to S_{1j}S_1C_j | S_{1j}S_{01}S_{2j}L_jS_2(1 \le j \le m)$
各々の偶数 k ($2 \le k \le n-4$) に対して：
$S_kC_j \to S_{kj}C_jS_k(1 \le j \le m)$
$S_kR_j \to S_{kj}S_{0k}S_{(k+1)j}S_{k+1}R_j(1 \le j \le m)$
$C_jS_{k+1} \to S_{(k+1)j}S_{k+1}C_j(1 \le j \le m)$
$L_jS_{k+1} \to S_{(k+1)j}S_{0(k+1)}S_{(k+2)j}L_jS_{k+2}(1 \le j \le m)$,
k=n-2 に対して：
$S_kC_j \to S_{kj}C_jS_k(1 \le j \le m)$
$S_kR_j \to S_{kj}S_{0k}S_{(k+1)j}S_{k+1}S_{(k+2)j}(1 \le j \le m)$
$C_jS_{k+1} \to S_{(k+1)j}S_{k+1}S_{(k+2)j}(1 \le j \le m)$
$L_jS_{k+1} \to S_{(k+1)j}S_{0(k+1)}S_{(k+2)j}(1 \le j \le m)$
$D_k \in \{L_k, R_k, C_k \mid 1 \le k \le m\}$ に対して,
$D_kS_{ij} \to S_{ij}D_k($ i=0, $1 \le j \le n-1$, または, $1 \le i \le n$, $1 \le j \le m)$.
(4.1.5) $n \ge 5$ かつ奇数のとき,
$S \to S_{01}S_{02} \cdots S_{0(n-1)} | S_{1j}S_{01}S_{2j}S_{02} \cdots S_{0(n-1)}S_{nj}(1 \le j \le m)$
$S \to S_{1j}S_1R_j(1 \le j \le m)$
$S_1 \to S_{1j}S_1C_j | S_{1j}S_{01}S_{2j}L_jS_2(1 \le j \le m)$
各々の偶数 k ($2 \le k \le n-5$) に対して：
$S_kC_j \to S_{kj}C_jS_k(1 \le j \le m)$
$S_kR_j \to S_{kj}S_{0k}S_{(k+1)j}S_{k+1}R_j(1 \le j \le m)$
$C_jS_{k+1} \to S_{(k+1)j}S_{k+1}C_j(1 \le j \le m)$
$L_jS_{k+1} \to S_{(k+1)j}S_{0(k+1)}S_{(k+2)j}L_jS_{(k+2)}(1 \le j \le m)$,
k=n-3 に対して：
$S_kC_j \to S_{kj}C_jS_k(1 \le j \le m)$
$S_kR_j \to S_{kj}S_{0k}S_{(k+1)j}S_{k+1}R_j(1 \le j \le m)$
$C_jS_{k+1} \to S_{(k+1)j}S_{k+1}C_j(1 \le j \le m)$
$L_jS_{k+1} \to S_{(k+1)j}S_{0(k+1)}S_{(k+2)j}S_{k+2}S_{(k+3)j}(1 \le j \le m)$
$S_{k+2}C_j \to S_{(k+2)j}S_{k+2}S_{(k+3)j}(1 \le j \le m)$
$S_{k+2}R_j \to S_{(k+2)j}S_{0(k+2)}S_{(k+3)j}(1 \le j \le m)$
$D_k \in \{L_k, R_k, C_k \mid 1 \le k \le m\}$ に対して,
$S_{ij}D_k \to D_kS_{ij}($ i=0, $1 \le j \le n-1$, または, $1 \le i \le n$, $1 \le j \le m)$
$D_kS_{ij} \to S_{ij}D_k($ i=0, $1 \le j \le n-1$, または, $1 \le i \le n$, $1 \le j \le m)$.

Gが$|E|$を生成することを確かめるためいくつかの例を考える．任意の$w \in |E|$を考える．ある$i \geq 0$，$1 \leq j_1, \cdots, j_i \leq m$に対して，$w \in L_1|E_{o1}|L_2\ |E_{o2}| \cdots |E_{on-1}|L_n$である．ここで，$L_k$は定義2.2のように定義される．

もし$i=0$であれば，$S \Rightarrow S_{o1}S_{o2} \cdots S_{on-1} \stackrel{*}{\Rightarrow} w.$

もし$i=1$であれば，$S \Rightarrow S_{1j_1}S_{o1}S_{2j_1}S_{o2} \cdots S_{on-1}S_{nj_1} \stackrel{*}{\Rightarrow} w.$

もし$i=2$かつ$n=4$であれば，

$S \Rightarrow S_{1j_1}S_1R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{o1}S_{2j_2}L_{j_2}S_2R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{o1}S_{2j_2}L_{j_2}S_{2j_1}S_{o2}S_{3j_1}S_3S_{4j_1}$

$\stackrel{*}{\Rightarrow} S_{1j_1}S_{1j_2}S_{o1}S_{2j_2}S_{2j_1}S_{o2}S_{3j_1}L_{j_2}S_3S_{4j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{o1}S_{2j_2}S_{2j_1}S_{o2}S_{3j_1}S_{3j_2}S_{o3}S_{4j_2}S_{4j_1}$

$\stackrel{*}{\Rightarrow} w.$

もし$i=3$かつ$n=5$であれば，

$S \Rightarrow S_{1j_1}S_1R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_1C_{j_2}R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}L_{j_3}S_2C_{j_2}R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}L_{j_3}S_{2j_2}C_{j_2}S_2R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}L_{j_3}S_{2j_2}C_{j_2}S_{2j_1}S_{o2}S_{3j_1}S_3R_{j_1}$

$\stackrel{*}{\Rightarrow} S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}S_{2j_2}S_{2j_1}S_{o2}S_{3j_1}L_{j_3}C_{j_2}S_3R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}S_{2j_2}S_{2j_1}S_{o2}S_{3j_1}L_{j_3}S_{3j_2}S_3C_{j_2}R_{j_1}$

$\Rightarrow S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}S_{2j_2}S_{2j_1}S_{o2}S_{3j_1}S_{3j_2}L_{j_3}S_3C_{j_2}R_{j_1}$

$W=S_{1j_1}S_{1j_2}S_{1j_3}S_{o1}S_{2j_3}S_{2j_2}S_{2j_1}S_{o2}S_{3j_1}S_{3j_2}$とおく，

$\Rightarrow WS_{3j_3}S_{o3}S_{4j_3}S_4S_{5j_3}C_{j_2}R_{j_1}$

$\stackrel{*}{\Rightarrow} WS_{3j_3}S_{o3}S_{4j_3}S_4C_{j_2}R_{j_1}S_{5j_3}$

$\Rightarrow WS_{3j_3}S_{o3}S_{4j_3}S_{4j_2}S_4S_{5j_2}R_{j_1}S_{5j_3}$

$\stackrel{*}{\Rightarrow} WS_{3j_3}S_{o3}S_{4j_3}S_{4j_2}S_4R_{j_1}S_{5j_2}S_{5j_3}$

$\Rightarrow WS_{3j_3}S_{o3}S_{4j_3}S_{4j_2}S_{4j_1}S_{o4}S_{5j_1}S_{5j_2}S_{5j_3}$

$\stackrel{*}{\Rightarrow} w$，等々．

(4.2) あるi, jに対して，$\lambda \in |E_{ij}|$である場合

まず，次の２つの集合を定義する：

$$Z=\bigcup_{i=1}^{n-1}\{S_{oi}\} \cup \left(\bigcup_{i=1}^{n}\bigcup_{j=1}^{m}\{S_{ij}\}\right);$$

$A=\{S_{01}S_{02}\cdots S_{0n-1}\}\cup\{W_1S_{01}W_2S_{02}\cdots S_{0n-1}W_n \mid$ ある $i\geq 1$, $1\leq j_1,\cdots,j_i\leq m$ に対して，(1) k が奇数なら $W_k=S_{kj_1}S_{kj_2}\cdots S_{kj_i}$, (2) その他は $W_k=S_{kj_i}S_{kj_{i-1}}\cdots S_{kj_1}\}$.

(4.1) で，われわれは $\{W\in Z^+ \mid S \overset{*}{\underset{G}{\Rightarrow}} W\}=A$ であることを見た.

任意の $S_{ij}\in X$ に対して，$L(S_{ij})$ は $|E_{ij}|$ を示す.

各 $W=a_1\cdots a_p\in Z^+$ $(p\geq 1,\ a_i\in Z)$ に対して，W_λ を次のように定義する：

$W_\lambda=\{W'\in Z^* \mid W'$ は $\lambda\in\bigcap_{j=1}^{q} L(a_{i_j})$ である $a_{i_1}\cdots,a_{i_q}$ $(q\geq 0)$ のいくらかの削除によって得られる$\}$.

$\lambda\in W_\lambda$ である必要十分条件はすべての $1\leq i\leq p$ に対して，$\lambda\in L(a_i)$ であることに注意する. 任意の $B\subset Z^*$ に対し，B_λ を次のように定義する：

$B_\lambda=\{W'\in Z^* \mid$ ある $W\in B$ に対して，$W'\in W_\lambda\}$.

任意の $j\geq i\geq 0$, $B\subset Z^*$ に対して，$B_{(\lambda,i,j)}$ を次のように定義する：

$B_{(\lambda,i,j)}=\{W\in B_\lambda \mid i\leq l(W)\leq j\}$.

任意の $1\leq i\leq n$ に対して，Y_i を次のように定義する：

$Y_i=\{j \mid 1\leq j\leq m$ かつ $\lambda\in|E_{ij}E_{(i+1)j}\cdots E_{nj}|\}$.

$|E|$ を生成する文法 $G_0=(V_0,\Sigma,P_0,S)$ を構成する方法を以下に与える：詳細のある部分は文脈により明かであり，省略する. 上記より，P_0 は

$\{W\in Z^* \mid S \overset{*}{\underset{G_0}{\Rightarrow}} W\}=A_\lambda$ でなければならない.

(4.2.1) $V_0=V\cup\{S_0\}\cup\{S_i',L_i',T_i,R_i',L_{ij},R_{ij},L_{ij}',R_{ij}',L_{ijp}',L_{ijpq}',R_{ijp}',R_{ijpq}' \mid 1\leq i,j,p,q\leq m\}\cup V_0'$;

$P_0=P\cup P_0'$.

ここで，V_0', P_0' は下記の議論から類推されるべきである. 任意の $\alpha\to\beta\in P_0$ に対しで，$l(\alpha)\leq l(\beta)$ であるから各々の $1\leq i,j\leq m$ に対して，(4.1) の C_iR_j と L_iC_j に相当する変数 R_{ij} と L_{ij} を用いる.

証明の残りの部分では，$n\geq 6$ かつ偶数の場合を考える.（その他の場合も議論は同様である）.

まず，P_0 に対して次の付加された規則が必要である：

(4.2.2) $S\to W$（各 $W\in A_{(\lambda,1,4n)}$ に対して）;

$S_0\to S_{1i}|S_{1i}S_0$（$i\in Y_2$ であれば）;

$1\leq i,j\leq m$ に対して，

$S\to S_{1i}S_{1j}S_1R_{ji}|S_0S_{1i}S_{1j}S_1R_{ji}|S_{1i}S_0S_{1j}S_1R_{ji}|S_{1i}S_{1j}S_0S_1R_{ji}$
$|S_0S_{1i}S_0S_{1j}S_1R_{ji}|S_0S_{1i}S_{1j}S_0S_1R_{ji}|S_{1i}S_0S_{1j}S_0S_1R_{ji}|S_0S_{1i}S_0S_{1j}S_0S_1R_{ji}$;

$S_1\to S_{1i}S_{1j}S_{01}S_{2j}S_{2i}L_{ji}S_2|S_{1i}S_0S_{1j}S_{01}S_{2j}S_{2i}L_{ji}S_2$
$|S_{1i}S_{1j}S_0S_{01}S_{2j}S_{2i}L_{ji}S_2|S_{1i}S_0S_{1j}S_0S_{01}S_{2j}S_{2i}L_{ji}S_2$;

$1\leq i\leq m$ に対して，

$S_1\to S_{1i}S_1C_i|S_0S_{1i}S_1C_i|S_{1i}S_0S_1C_i|S_0S_{1i}S_0S_1C_i$.

以下の規則の集合より，$S_{1k_1}S_{1k_2}\cdots S_{1k_q}\in|S_{1j_1}\cdots S_{1j_p}|_\lambda$ かつすべての

$j_r \notin \{k_1, \cdots, k_q\}$に対して$j_r \in Y_2$である任意の
$W = S_{1j_1} S_{1j_2} \cdots S_{1j_p} S_{01} S_{2k_1} S_{2k_2} L_{k_1 k_2} C_{k_3} \cdots C_{k_{q-2}} R_{k_{q-1} k_q}$に対して,

$S \overset{*}{\underset{G_0}{\Rightarrow}} W$が成立する.

次に, k ($k \geq 2$) 番目の段階を考える.
kを偶数かつ$k \leq n-2$ とする (kが奇数のときも同様に議論される).
まず, 次の(4.2.3)が必要である:
(4.2.3) $S_k C_i \to S_{ki} S_k$($i \in Y_{k+1}$であれば); $S_k C_i \to C_i S_k$ ($\lambda \in |E_{ki}|$であれば).
(4.2.3.*) さらに, 次のk番目の段階の場合を考える:

$$S \overset{*}{\Rightarrow} W S_{0k-1} S_{kj_1} S_{kj_2} L_{j_1 j_2} S_k C_{j_3} \cdots C_{j_{u-2}} R_{j_{u-1} j_u}$$

$$\overset{*}{\Rightarrow} W S_{0k-1} W_1 S_{0k} S_{(k+1)k_1} S_{(k+1)k_2} L_{k_1 k_2} S_{k+1} C_{k_3} \cdots C_{k_{v-2}} R_{k_{v-1} k_v}$$

が成立し, かつ$v < u$である. この場合次の(4.2.3.2)が必要である:
(4.2.3.2) $S_k \to S_k'$; $S_k' \to S_k$;
もし$i, q \in Y_{k+1}$であれば, $1 \leq j,p,r,t \leq m$に対して,
$L_{ij} S_k' C_p \to S_{kp} L_{jp} S_k'$; $S_k C_j \to S_{kj} R_j' S_k$;
$R_j' S_k C_q \to S_{kq} R_j' S_k$; $R_j' S_k C_p \to S_{kp} R_{jp}' S_k$;
$R_{jp}' S_k C_q \to S_{kq} R_{jp}' S_k$; $R_{jp}' S_k R_{rt} \to S_{kr} S_{kt} S_{0k} S_{(k+1)t} S_{(k+1)r} S_{k+1} R_{jp}$.

上の場合, 最後にR_{rt}はR_{jp}に置き換える. その他, R_{rt}がR_{pr}に置き換えるような可能性もありうる: しかし議論は同様である. $k=n-2$に対して, 次の(4.2.4)が必要である:
(4.2.4) $S_k C_i \to S_{ki} S_k$(もし$i \in Y_{k+1}$であれば);
$S_k C_i \to C_i S_k$(もし$\lambda \in |E_{ki}|$であれば);
$S_k R_{ij} \to S_{ki} S_{kj} S_{0k} S_{(k+1)j} S_{(k+1)i} S_{k+1}$;
$C_i S_{k+1} \to S_{(k+1)i} S_{k+1} S_{(k+2)i}$;
$L_{ij} S_{k+1} \to S_{(k+1)j} S_{(k+1)i} S_{0(k+1)} S_{(k+2)i} S_{(k+2)j}$.

上の各々の生成規則$\alpha \to \beta$において, $l(W) \geq l(\alpha)$である各々の$W \in \beta_\lambda$に対して, 規則$\alpha \to W$も必要である.

最後に入る前, 次の2つの概念を定義する:
$2 \leq i \leq n$, $1 \leq j \leq 3$, $1 \leq k_1, \cdots, k_j \leq m$に対して,
$X_{i(k_1, \ldots, k_j)} = W_i S_{0i} W_{i+1} \cdots S_{0n-1} W_n$である.
ここで, もしpが奇数であれば, $W_p = S_{pk_1} S_{pk_2} \cdots S_{pk_j}$である,
その他は, $W_p = S_{pk_j} S_{pk_{j-1}} \cdots S_{pk_1}$である;
$X_i = \{ S_{0i} S_{0i+1} \cdots S_{0n-1} \} \cup \{ W \in X_{i(k_1, \ldots, k_j)} \mid 1 \leq j \leq 3$かつ$1 \leq k_1, \cdots, k_j \leq m \}$.

さて, 最後に次の場合を考える:
あるk ($1 \leq k \leq n-1$), $W_0, W_1, W_2 \in Z^*$に対し, $l(W_1) \geq 4$, $W_0 S_{0k-1} W_1 S_{0k} W_2 \in A_\lambda$, $W_2 \in (X_{k+1})_\lambda$が成立する. ここで, もし$k=1$であれば, $W_0 S_{0k-1} = \lambda$である.
次の場合を考える:

ある $1 \leq j_1, j_2, j_3 \leq m$ に対して,k は偶数かつ, $\leq n-4$ かつ $W_2 \in X_{k+1(j_1,j_2,j_3)}$（その他の場合は同様に扱える）. 次の規則が必要である：

(4.2.5) $L_{ij}S_{k-1} \rightarrow S_{k-1j}S_{k-1i}S_{0k-1}S_{ki}L_{ijj}'T_k$;

もし $i, u \in Y_{k+1}$ であれば, $1 \leq j, p, q, r, t \leq m$ に対して,

$L_{ij}S_{k-1} \rightarrow S_{k-1j}S_{k-1i}S_{0k-1}S_{ki}L_{jj}'T_k$;

$L_{iq}'T_kC_r \rightarrow S_{kq}L_{rr}'T_k$ （もし $q \in Y_{k+1}$ であれば） ;

$L_{pq}'T_kC_i \rightarrow S_{kq}L_{pi}'T_k$; $L_{pq}'T_kC_r \rightarrow S_{kq}L_{prr}'T_k$;

$L_{pqr}'T_kC_i \rightarrow S_{kr}L_{pqi}'T_k | S_{kr}L_{pqii}'T_k$; $L_{pqr}'T_kC_t \rightarrow S_{kr}L_{pqtt}'T_k$;

$L_{pqrt}'T_kC_i \rightarrow S_{kt}L_{pqri}'T_k$;

$L_{pqrt}'T_kR_{iu} \rightarrow S_{kt}S_{ki}S_{ku}S_{0k}W$ （各 $W \in (X_{k+1(p,q,r)})_\lambda$ に対して）.

ここで, 例えば, L_{pq}' は次のステップで S_{kq} が置かれ, そして S_{kp} は最後のステップのため, 覚えておくべき記号であることを示す非終端記号である, 等々.

上の場合 $W \in (X_{k+1(p,q,r)})_\lambda$ である. $W \in (X_{k+1(q,r,i)})_\lambda \cup (X_{k+1(r,i,u)})_\lambda$ であるような可能性もありうる. しかし議論は同様である. このようにして, ある i, j に対して, $\lambda \in |E_{ij}|$ である場合も, $|E|$ を生成する文脈依存文法 G_0 を実際に構成することが可能であることが証明される. □

［例3. 1］

定理3. 2の(4. 1)の証明で, P' は次の生成規則の集合からなる.

(1) n=7, m=3 の場合

$S \rightarrow S_{01}S_{02} \cdots S_{06} | S_{1j}S_{01}S_{2j}S_{02} \cdots S_{06}S_{7j} (1 \leq j \leq 3)$

$S \rightarrow S_{1j}S_1R_j (1 \leq j \leq 3)$

$S_1 \rightarrow S_{1j}S_1C_j | S_{1j}S_{01}S_{2j}L_jS_2 (1 \leq j \leq 3)$

$S_2C_j \rightarrow S_{2j}C_jS_2 (1 \leq j \leq 3)$

$S_2R_j \rightarrow S_{2j}S_{02}S_{3j}S_3R_j (1 \leq j \leq 3)$

$C_jS_3 \rightarrow S_{3j}S_3C_j (1 \leq j \leq 3)$

$L_jS_3 \rightarrow S_{3j}S_{03}S_{4j}L_jS_4 (1 \leq j \leq 3)$.

$S_4C_j \rightarrow S_{4j}C_jS_4 (1 \leq j \leq 3)$

$S_4R_j \rightarrow S_{4j}S_{04}S_{5j}S_5R_j (1 \leq j \leq 3)$

$C_jS_5 \rightarrow S_{5j}S_5C_j (1 \leq j \leq 3)$

$L_jS_5 \rightarrow S_{5j}S_{05}S_{6j}S_6S_{7j} (1 \leq j \leq 3)$

$S_6C_j \rightarrow S_{6j}S_6S_{7j} (1 \leq j \leq 3)$

$S_6R_j \rightarrow S_{6j}S_{06}S_{7j} (1 \leq j \leq 3)$

$D_k \in \{L_k, R_k, C_k \mid 1 \leq k \leq 3\}$ に対して,

$S_{ij}D_k \rightarrow D_kS_{ij}$ (i=0, $1 \leq j \leq 6$, または, $1 \leq i \leq 7$, $1 \leq j \leq 3$)

$D_kS_{ij} \rightarrow S_{ij}D_k$ (i=0, $1 \leq j \leq 6$, または, $1 \leq i \leq 7$, $1 \leq j \leq 3$).

(2) n=8, m=3 の場合

$S \rightarrow S_{01}S_{02} \cdots S_{07} | S_{1j}S_{01}S_{2j}S_{02} \cdots S_{07}S_{8j} (1 \leq j \leq 3)$

$S \rightarrow S_{1j}S_1R_j (1 \leq j \leq 3)$

$S_1 \to S_{1j}S_1C_j | S_{1j}S_{01}S_{2j}L_jS_2 (1 \le j \le 3)$
$S_2C_j \to S_{2j}C_jS_2 (1 \le j \le 3)$
$S_2R_j \to S_{2j}S_{02}S_{3j}S_3R_j (1 \le j \le 3)$
$C_jS_3 \to S_{3j}S_3C_j (1 \le j \le 3)$
$L_jS_3 \to S_{3j}S_{03}S_{4j}L_jS_4 (1 \le j \le 3)$
$S_4C_j \to S_{4j}C_jS_4 (1 \le j \le 3)$
$S_4R_j \to S_{4j}S_{04}S_{5j}S_5R_j (1 \le j \le 3)$
$C_jS_5 \to S_{5j}S_5C_j (1 \le j \le 3)$
$L_jS_5 \to S_{5j}S_{05}S_{6j}L_jS_6 (1 \le j \le 3)$
$S_6C_j \to S_{6j}C_jS_6 (1 \le j \le 3)$
$S_6R_j \to S_{6j}S_{06}S_{7j}S_7S_{8j} (1 \le j \le 3)$
$C_jS_7 \to S_{7j}S_7S_{8j} (1 \le j \le 3)$
$L_jS_7 \to S_{7j}S_{07}S_{8j} (1 \le j \le 3)$

$D_k \in \{L_k, R_k, C_k \mid 1 \le k \le 3\}$ に対して,
$D_kS_{ij} \to S_{ij}D_k (i=0,\ 1 \le j \le 7,$ または, $1 \le i \le 8,\ 1 \le j \le 3)$.

[定理3. 3]

任意の $n \ge 3$ に対して,

(1) $ERL(n, \Sigma) - CFL(\Sigma) \ne \phi$;

(2) $LCFL(\Sigma) - ERL(n, \Sigma) \ne \phi$.

ここで, $\#\Sigma \ge 3$.

(証明) $\Sigma = \{a, b, c\}$ とする.

(1) $\{a^nb^na^n \mid n \ge 0\} \in ERL(n, \Sigma) - CFL(\Sigma)$.

(2) 次の線形文法 $G=(V, \Sigma, P, S)$ を考える.

ここで, $\Sigma=\{a, b, c\}$, $V=\{S, A\}$,

$P: S \to aba|aSa|cAc$,

$A \to bAb|cSc$ である.

任意の $n \ge 3$ に対して, $L(G) \notin ERL(n, \Sigma)$ であることを証明すればよい. 証明は [7] の定理6. 1の(5)の証明と同じ方法で行う. $L(G) \in ERL(n,\Sigma)$ と仮定する. そのとき $L(G)$ を示す $E \in ERE(n, \Sigma)$ が存在する. まず, 次の命題3.1を示す.

[命題3. 1]

(1) 任意の $w \in L(G)$ に対して, $w=vabav^R$ であるような $v \in \Sigma^*$ が存在する;

(2) $L(G) \cap \Sigma^*aba\Sigma^*aba\Sigma^*=L(G) \cap \Sigma^*abb\Sigma^*=L(G) \cap \Sigma^*bba\Sigma^* = \phi$;

(3) 任意の $u, v, w, x, y \in \Sigma^*$ に対して, もしすべての $i \ge 0$ に対して $ux^ivy^iw \in L(G)$ であれば, $l(x)=l(y)$ である.

(4) 任意の $k \ge 2$, $u_1, \cdots u_{k+1} \in \Sigma^*$, $L_1, \cdots, L_k \subset \Sigma^*$ に対して, もし各々の L_j が空でなく, かつすべての $i \ge 0$ に対して $u_1(L_1)^iu_2(L_2)^i \cdots u_k(L_k)^iu_{k+1} \subset L(G)$ であれば, 各々の L_j は有限である.

(命題の証明) (1) - (3) は[7]の定理6. 1で証明された.

(4). 条件を仮定する. ある L_j $(1 \le j \le k)$ は無限であると仮定する. L_1 が無限であると仮定する (その他の場合も同様). そのとき, すべての j に対して, $v_j \in$

L_j が存在する．従って，$u_1(L_1)u_2v_2 \cdots u_kv_ku_{k+1} \subset L(G)$ である．L_1 が無限であるから，$l(t) \geq 2l(u_1u_2v_2 \cdots u_kv_kv_{k+1}+3)$ であるような $t \in L_1$ が存在する．(1) より，ある t_0, $t_1 \in \Sigma^*$ に対して，$t=t_0abat_1$ である．このとき，$u_1(t)^2u_2(v_2)^2 \cdots u_k(v_k)^2u_{k+1} \in L \cap \Sigma^*aba\Sigma^*aba\Sigma^*$ である．これは(2)に矛盾する． □

定理3.3の証明（つづき）．pをEに現われる$\Sigma \cup \{\lambda\}$に属するすべての記号の数とする．次の語wを考える：

$w=(a^{p+1}cb^{p+1}c)^{p+1}aba(cb^{p+1}ca^{p+1})^{p+1}$．明らかに$w \in L(G)$である．従って，wはEの中で綴られる．Eの位置iに，位置iに現われる各々の$x \in \Sigma$を結びつける．ここで，iは正整数でかつxは左からEのi番目の位置に現われる記号を示す．従って，wに対して，ペアの系列$((x_1,i_1), (x_2,i_2), \cdots,(x_p,i_p))$が存在する．ここで$w=x_1x_2 \cdots x_p$，$x_1= \cdots =x_{p+1}=a$, $x_{p+2}=c$, $x_{p+3}= \cdots =x_{2p+3}=b, \cdots$ 等，そして各々のi_jはEにおけるx_jの位置である．pの定義によって$1 \leq j_{11} < j_{11}' \leq p+1$かつ$i_{j_{11}}=i_{j_{11}'}$であるような$j_{11}, j_{11}'$が存在する．命題3.1の(3)によって(H)*の形のEの部分表現は存在しないことに注意する．従って$i_{j_{11}}$番目の位置は次のように$m^*(2 \leq m \leq n)$の中に現われる：

$$\cdots\left[\cdots,\cdots,\cdots,\left[\begin{array}{c}\vdots\\ a^k\\ \vdots\end{array}\right],\cdots,\cdots\right]^{m*}$$

同様に$p+3 \leq j_{21} < j_{21}' \leq 2p+3$かつ$i_{j_{21}}=i_{j_{21}'}$であるような$j_{21}, j_{21}'$が存在する．従って$i_{j_{21}}$番目の位置は次のように$r^*(2 \leq r \leq n)$の中に現われる：

$$\cdots\left[\cdots,\cdots,\left[\begin{array}{c}\vdots\\ b^s\\ \vdots\end{array}\right],\cdots,\cdots\right]^{r*}$$

命題3.1の(2)より $\left[\begin{array}{c}\vdots\\ a^k\\ \vdots\end{array}\right]$ と $\left[\begin{array}{c}\vdots\\ b^s\\ \vdots\end{array}\right]$ は明かに異なる．wとEを通じて整数の系列

$j_{11}, \cdots, j_{1p+1}, j_{21}, \cdots, j_{2p+1}$を得た．pの選択によって$1 \leq e < f \leq p+1$かつ$i_{j_{1e}}=i_{j_{1f}}$であるような整数e, fが存在する．従って，Eは次のような部分表現を持つ；

$$\left[\cdots,\left[\cdots\left(\cdots,\cdots,\begin{array}{c}\vdots\\ \cdots\\ \vdots\end{array},\cdots\right)^{u*}\cdots\right],\cdots\right]^{v*}$$

これは，命題3.1の(4)に矛盾する． □

参考文献

[1] S. Ginsburg, The Mathematical Theory of Context-free Languages, (McGraw-Hill, 1966).

[2] S. Ginsburg and H.G. Rice, Two families of languages related to ALGOL, J.ACM, 9, pp.350-371 (1962).

[3] S. Ginsburg and E.H. Spanier, Bounded ALGOL-like languages, Trans. AM. Math.

Soc., 113, pp.333-368 (1964).
[4] S.A. Greibach, The unsolvability of the recognition of linear context-free languages, JACM, 13, 4, pp.582-587 (1966).
[5] J. Gruska, A characterization of context-free languages, J. Comput. System Sci., 5, pp.353-364 (1971).
[6] M. A. Harrison, Introduction to Formal Language Theory, (Addison-Wesley, 1978).
[7] K. Hashiguchi and H. Yoo, Extended regular expressions of star degree at most two, to appear in Theoret. Comput. Sci.
[8] K. Hashiguchi and H. Yoo, The infinite 2-star height hierarchy of extended regular languages of star degree at most two, submitted to Theoret. Comput. Sci.
[9] J. E. Hopcroft and J. D. Ullman, Introduction to Automaton Theory, Languages, and Computation, (Addison-Wesley, 1979).
[10] E. Shamir, On sequential languages, Z. Phonetik, Sprach. & Kommunikationsforschung, 18, pp.61-69 (1965).
[11] 山崎, 文脈自由言語のある部分言語の包含関係に関する一考察, 信学論(D), J70-D, 1, pp.1-9(87/1).
[12] M. K. Yntema, Cap expressions for context-free languages, Inform. Control., 18, pp.311-318 (1971).
[13] H. Yoo and K. Hashiguchi, Extended automata-like regular expressions of star degree at most (2,1), to appear in Theoret. Comput. Sci.